npg

Local adaptation in the rock pocket mouse (*Chaetodipus intermedius*): natural selection and phylogenetic history of populations

HE Hoekstra, JG Krenz[1] and MW Nachman
Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721, USA

Elucidating the causes of population divergence is a central goal of evolutionary biology. Rock pocket mice, *Chaeotdipus intermedius*, are an ideal system in which to study intraspecific phenotypic divergence because of the extensive color variation observed within this species. Here, we investigate whether phenotypic variation in color is correlated with local environmental conditions or with phylogenetic history. First, we quantified variation in pelage color ($n = 107$ mice) and habitat color ($n = 51$ rocks) using a spectrophotometer, and showed that there was a correlation between pelage color and habitat color across 14 sampled populations ($R^2 = 0.43$). Analyses of mtDNA sequences from these same individuals revealed strong population structure in this species across its range, where most variation (63%) was partitioned between five geographic regions. Using Mantel tests, we show that there is no correlation between color variation and mtDNA phylogeny, suggesting that pelage coloration has evolved rapidly. At a finer geographical scale, high levels of gene flow between neighboring melanic and light populations suggest the selection acting on color must be quite strong to maintain habitat-specific phenotypic distributions. Finally, we raise the possibility that, in some cases, migration between populations of pocket mice inhabiting different lava flows may be responsible for similar melanic phenotypes in different populations. Together, the results suggest that color variation can evolve very rapidly over small geographic scales and that gene flow can both hinder and promote local adaptation.
Heredity (2005) **94,** 217–228. doi:10.1038/sj.hdy.6800600
Published online 3 November 2004

Introduction

A central goal of developmental and evolutionary biology is to explain the morphological diversity observed across species. As differences among species must initially occur as intraspecific polymorphism, understanding the causes of intraspecific variation can provide information about the origin of species-level differences. One ongoing debate concerns the relative roles of deterministic evolutionary processes and historical contingency in shaping the outcome of evolution (Travisano *et al*, 1995). It has long been appreciated that strong selection can lead to local adaptation, provided that it is not swamped by gene flow (Haldane, 1948; Slatkin, 1985; Lenormand, 2002). On the other hand, populations may sometimes be constrained by their evolutionary history. For example, young populations in new environments may not have had time to adapt to

local conditions. In other cases, pleiotropic effects of otherwise beneficial mutations may limit their spread. Here, we focus on the extent to which phenotypic variation is correlated with local environmental conditions versus phylogenetic history.

The rock pocket mouse, *Chaetodipus intermedius*, provides an excellent system to study geographic variation in phenotype within a single species, and allows us to explore this variation in light of the underlying genetic structure of this species. In particular, we have been interested in the evolution of color differences in response to local environmental conditions. *C. intermedius* lives exclusively in rocky habitat across the southwestern deserts, and thus *C. intermedius* habitat is largely discontinuous through most of its range. Historically, *C. intermedius* comprises 10 subspecies (Benson, 1933; Dice and Blossom, 1937; Weckerly, 1983). Several subspecies have been described based on dramatic color differences on small isolated lava flows (Benson, 1933); these mice have extremely dark coats and uniformly melanic hairs. Non-lava-dwelling populations also show variation in coat color, which often closely resembles the substrate color on which the mice live (Benson, 1933; Dice and Blossom, 1937).

Previous work identified the genetic basis of melanism in a single population of mice inhabiting the Pinacate lava flow in southern Arizona (Nachman *et al*, 2003). Four amino-acid changes in the coding region of the melanocortin-1 receptor (*Mc1r*) are perfectly associated

Correspondence: HE Hoekstra, Current address: Ecology, Behavior and Evolution Section, Division of Biological Sciences, University of California at San Diego, 9500 Gilman Drive, MC0116, Muir Biology Bldg, Rm 4256, La Jolla, CA 92093-0116, USA.
E-mail: hoekstra@ucsd.edu
[1]Current address: JG Krenz, United States Department of Agriculture, Agricultural Research Service-Aquaculture Genetics, Oregon State University-Hatfield Marine Science Center, 2030 SE Marine Science Dr, Newport, OR 97365, USA.

with a coat color polymorphism in this population. A focused study on the Pinacate region suggested that strong selection is maintaining *Mc1r* allele and coat color frequencies across short geographic distances in the face of high countervailing gene flow (Hoekstra *et al*, 2004). Interestingly, mutations in the *Mc1r* gene are not involved in additional melanic populations in New Mexico (Hoekstra and Nachman, 2003). While genes unlinked to those that underlie the phenotype of interest cannot provide a direct test of whether melanism has arisen independently in New Mexico, they can provide information about population structure, levels of gene flow and timing of colonization, which may have implications for the evolution of melanic populations.

More generally, the extensive phenotypic variation observed in *C. intermedius* raises several interesting questions. First, are morphological differences reflected in the genetic structure of the species? How much gene flow occurs between populations which differ in phenotype? Finally, what is the time scale over which this phenotypic variation has evolved? To address these questions, we have quantified phenotypic variation in coat color and local substrate color across 14 populations of *C. intermedius*. We used a spectrophotometer to quantify color variation among populations and sequenced two mtDNA genes to characterize the population structure and extent of gene flow between populations relative to phenotypic differentiation.

Methods

Phenotypic variation

Sampling: A total of 107 rock pocket mice were collected using Sherman live traps from 14 localities across the species range in Arizona, New Mexico and northern Mexico (Figure 1, Table 1). To quantify variation in local habitat, rocks were collected in areas neighboring the trap-lines at each site. Localities were chosen to maximize phenotypic and environmental variation, rather than to cover the species range. In addition, several localities represent paired sampling sites, where substrate color differed dramatically over short geographical distances. Liver, kidney and spleen samples were taken from each individual and frozen at $-80°C$. Voucher specimens were prepared and deposited in the Zoological Collections of the Department of Ecology and Evolutionary Biology at the University of Arizona.

Phenotypic and environmental variation: The reflectance of both mouse coat color and corresponding rocks was measured as a percentage of a white standard using a USB2000 spectrophotometer (Ocean Optics) with a dual deuterium/halogen light source. A standard reflection probe with a 200μ receptor fiber was held at an uniform distance from the surface at $90°$ to capture both diffuse and spectral reflectance. Measurements were taken at 1518 points from 300 to 800 nm, and thus included the UV spectrum.
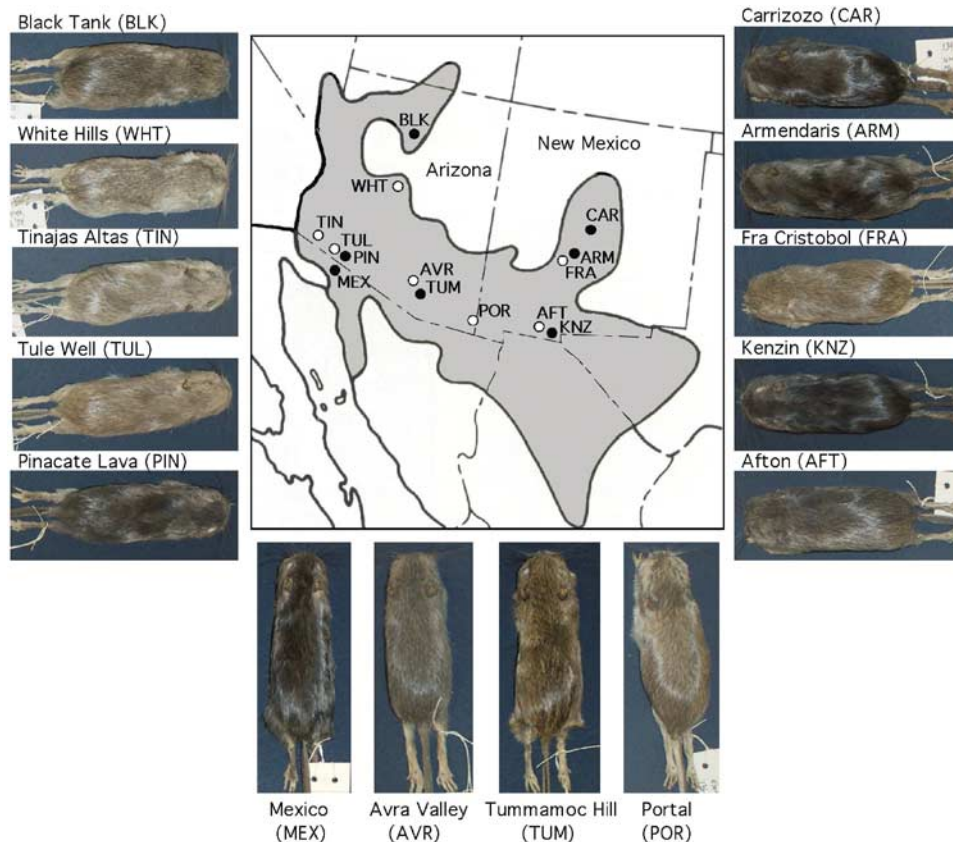
Figure 1 Phenotypic variation across the range of *C. intermedius* in the southwestern US. Photographs represent the typical dorsal coloration of individuals from each of 14 collecting locales indicated by circles. Filled circles represent lava flows and open circles are nonvolcanic rocky regions. The border between Arizona and New Mexico roughly represents the interface of the Sonoran and Chihuahuan deserts, respectively.

Table 1 Collecting localities

Locale		County	N	Latitude/longitude	Subspecies	Specimen numbers
Black Tank Lava	BLK	Coconino Co., AZ	3	35°35.4′/111°38.1′	C. i. crinitus	HEH 615-7
White Hills	WHT	Yavapai Co., AZ	10	34°39.3′/111°54.0′	C. i. umbrosus	HEH 618-29
Tinajas Altas	TIN	Yuma Co., AZ	9	32°18.8′/114°02.9′	C. i. phasma	HEH 635-9, 644-5, 648-9
Tule Mts.	TUL	Yuma Co., AZ	10	32°10.5′/113°40.2′	C. i. mearnsi	HEH 546-8, 550-3, 589, 592, 603
Pinacate Lava	PIN	Yuma Co., AZ	10	32°05.9′/113°28.3′	C. i. pinacate	HEH 534-42, MWN 1434
Pinacate Mexico	MEX	Sonora, MX	3	32°06.5′/113°22.5′	C. i. pinacate	HEH 512-4
Avra Valley	AVR	Pinal Co., AZ	10	32°23.8′/111°08.5′	C. i. intermedius	MWN 1358-67
Tumamoc Lava	TUM	Pima Co., AZ	3	32°13.0′/111°00.4′	C. i. nigrimontis	HEH 509-511
Portal	POR	Cochise Co., AZ	5	31°54.9′/109°03.7′	C. i. intermedius	MWN 1394-5, 1399-1401
Afton Hill	AFT	Dona Ana Co., NM	6	32°09.6′/107°10.4′	C. i. intermedius	HEH 515-20
Kenzin Lava	KNZ	Dona Ana Co., NM	10	32°03.4′/106°57.9′	C. i. rupestris	HEH 504-6, 522-4, 526, 557, 559, 687
Fra Cristobol Mts.	FRA	Sierra Co., NM	10	33°19.9′/107°07.1′	C. i. intermedius	HEH 571, 574-582
Armendaris Lava	ARM	Socorro Co., NM	9	33°31.0′/106°60.0′	(un-named)	HEH 521, 561-8
Carrizozo Lava	CAR	Lincoln Co., NM	9	33°41.5′/105°50.9′	C. i. ater	MWN 1337-45

In all, 107 nonmolting adults and 51 rocks were analyzed from all 14 collecting sites. For each animal, 10 measurements were made from the dorsal surface of the mouse and averaged to produce a general description of the dorsal coat color. Similarly, 10 measurements from the exposed rock surface were taken and averaged. Analysis followed that described in Hoekstra and Nachman (2003). Here, we report total reflectance (relative to a pure white standard) to characterize overall dorsal pelage and rock surface coloration.

Genetic variation

Genomic DNA was prepared using a tissue extraction kit (Qiagen DNeasy Tissue kit). MtDNA genes, COIII (783 bp) and ND3 (345 bp), were amplified from each individual. In some populations, the intergenic tRNA-Gly was included in the analysis, so the total genetic region analyzed was 1176 bp in length. Primer sequences and PCR conditions are reported in Hoekstra et al (2004). PCR products were cleaned using spin columns (Qiagen) and were sequenced on an ABI 3700. Sequences were assembled and aligned in Sequencher (GeneCodes) and checked by eye. Outgroup sequences were obtained from the sister species, C. penicillatus (GenBank AY259036) and C. baileyi (AY259035). Sequences have been deposited in Genbank (accession numbers AY694010-AY694094, see Appendix A1).

Summary statistics: The sequence alignment was imported into DnaSP v3.0 (Rozas and Rozas, 1999) to calculate intra- and interspecific genetic variability. In each population, the number of segregating sites and the number of unique haplotypes were counted. The average number of pairwise differences, π (Nei and Li, 1979), and diversity based on the number of segregating sites, θ (Waterson, 1975), were calculated. To check for deviations from neutral expectations in the frequency spectrum of polymorphisms, significance values were calculated for Tajima's D statistic (Tajima, 1989).

Phylogenetic analysis: To estimate the phylogenetic relationships among haplotypes and among populations, gene genealogies were constructed. Gene trees were generated in PAUP* v4.02 (Swofford, 1999) using neighbor-joining (NJ), maximum parsimony (MP) and maximum-likelihood (ML) algorithms. NJ trees were generated using the transition/transversion (ti/tv) rate and gamma distribution shape parameter (γ) estimated from Modeltest v3.06 (Posada and Crandall, 1998). Parsimony genealogies were generated with transitions and transversions weighted equally, and also with transversions given a weight two, five and 10 times more than that of transitions; the observed number of transitions ($n = 250$) in our data was 4.6 times greater than the observed number of transversions ($n = 55$). Heuristic searches were performed with stepwise addition for initial trees and the tree-bisection-reconnection method of branch swapping. The consensus parsimony tree was used as the starting tree in the estimation of ML trees. Hierarchical likelihood-ratio tests implemented in Modeltest were used to determine the best-fit model of nucleotide substitution to estimate phylogenetic trees and genetic distances. Topologies were explored using a branch-and-bound method. Confidence in the branching structure was assessed by performing 1000 bootstrap replicates.

Population structure and gene flow: To test for significant population structure among populations and biogeographical regions, analyses of molecular variance (AMOVA, Excoffier et al, 1992) were performed in ARLEQUIN (Schneider et al, 2000). Pairwise F_{ST} estimates were permuted 1000 times. A one-factor AMOVA was employed to assess the degree of population structure over all populations. Broad-scale patterns of regional diversity were examined by pooling samples within geographical regions identified as having a recent common history based on phylogenetic analyses. These pooled samples also correlated to known biogeographical regions. We clustered populations at three levels: (1) Sonoran versus Chihuahuan deserts (Sonoran = BLK + WHT + TIN + TUL + PIN + MEX + AVR + TUM and Chichuahuan = POR + AFT + KNZ + FRA + ARM + CAR), (2) five clusters based on fine-scale biogeographic regions (northern Arizona = BLK + WHT, southern Arizona = TIN + TUL + PIN + MEX, central Arizona = TUM + AVR, southern New Mexico = POR + AFT + KNZ, central New Mexico = FRA + ARM + CAR) and (3) northern Arizona versus all other populations (northern Arizona = BLK + WHT).

We tested for isolation by distance between populations of *C. intermedius* by estimating pairwise effective migration rates $Nm = [(1/F_{ST})-1]/2$ between all possible pairs of populations, where N is the female effective population size, m is the female migration rate and F_{ST} is a measure of population structure (Slatkin, 1993). We also calculated F_{ST} following the method suggested by Rousset (1997), which is a modification of Slatkin's F_{ST} using an isolation by distance approach, and results were consistent using both methods. When populations are in migration–drift equilibrium and isolated by distance, the effective migration rate should be negatively correlated with interpopulation geographic distance.

Correlation between phenotype, genotype and geography

To test for a significant correlation between phenotypic, genetic and geographical distance among populations of *C. intermedius*, we performed Mantel tests using ARLE-QUIN. Mantel procedures can test for an association between two matrices using randomization (Manly, 1986). Specifically, the parameter is compared to a distribution obtained when the matrix is repeatedly randomized, and the null hypothesis of no association is rejected when the parameter exceeds a given significance level. In this case, we can test whether patterns of color variation are significantly associated with geography (implying local adaptation) and/or genetic distances (the influence of shared evolutionary history). We also conducted partial Mantel tests, which hold one matrix constant and test for an association between the remaining two matrices. As the appropriateness of partial Mantel tests have recently been called into question, we use the results from the partial Mantel tests only to support results from the Mantel tests (Raufaste and Rousset, 2001; Castellano and Balletto, 2002; Rousset, 2002).

Population divergence and local adaptation

We employed a molecular clock to generate estimates of divergence times between clades. This clock is based on silent site divergence of mtDNA genes in rodents of approximately 2% per million years (Wilson *et al*, 1985). We compared average pairwise genetic differences between: (1) individuals from the northern Arizona populations and all other individuals and (2) individuals from the central Arizona populations and individuals from New Mexico. The latter comparison provides an age estimate for the colonization of New Mexico by pocket mice, and therefore a maximum age for the melanic populations on the three New Mexico lava flows.

To test the role of local adaptation in phenotypic evolution, we examined the extent of gene flow occurring between each lava population and the nearest population on light-colored rocks. We made the following population pairwise comparisons: ARM and FRA, KNZ and AFT, PIN and TUL, and TUM and AVR. For each of the four comparisons, we calculated F_{ST} and Nm using DnaSP. Similarly, we measured gene flow (F_{ST} and Nm) between the three lava-dwelling populations in New Mexico to explore the role of migration in generating melanic phenotypes on each of the lava flows. For all seven pairwise comparisons, we also used a coalescent-based approach implemented in the pro-gram MDIV (Nielsen and Wakeley, 2001). This Markov chain Monte Carlo (MCMC) method allows us to determine whether shared polymorphisms are the result of recurrent gene flow, recent common ancestry or both. We used a finite-sites mutation model (HKY) following Palsboll *et al* (2004).

Results

Phenotypic variation

A total of 107 mice and 51 rocks from 14 populations across the species range were included in the spectro-photometric measurements (Figure 2). The percent reflectance from the dorsal mouse coat ranged from 4.02 to 11.96%. The darkest six populations represented mice inhabiting lava flows, indicated by an asterisk in Figure 2.

Reflectance from the substrate was significantly higher than mouse coat reflectance, ranging from 9.18 to 41.8%. There is a clear difference between nonvolcanic rocks (eg granite, gneiss and limestone) and rocks of volcanic origins ($t = -7.43$; $P < 0.0001$). The mean percent reflectance for nonvolcanic rocks was 7.93% (SD = 0.31) and for volcanic rocks was 4.49% (SD = 0.35).

There was a significant positive correlation between mean rock color and mean mouse color among the 14 populations sampled ($R^2 = 0.429$, $P < 0.05$; Figure 2). However, there are two notable exceptions. First, one population from northern Arizona, Black Tank Lava (BLK), had relatively light-colored mice but the rock was very dark and volcanic in origin. Second, the rock substrate from the White Hills (WHT) in northern Arizona is extremely light-colored (the lightest reported here); however, reflectance from the corresponding mice was unremarkable.

Genetic variation

The mtDNA loci surveyed in this study exhibit high levels of polymorphism in terms of number of haplotypes, segregating sites and nucleotide diversity as measured by π and θ (Table 2). Across the species, there were 67 haplotypes and 136 polymorphic sites in the combined 1111 bp of the COIII and ND3 genes (Figure 3). The mean value of nucleotide diversity was also high, $\pi = 0.014$ and $\theta = 0.024$. These estimates are similar to nucleotide levels observed in other small mammals (Nachman *et al*, 1994).

Intrapopulation genetic variation: Between three and 10 mice were surveyed from each of 14 populations (Figure 1; Table 2). Highest levels of intrapopulation diversity were in the southern Arizona populations, including Tinajas Altas Mts (TIN; $\pi = 0.0071$), Cabeza Prieta Mts at Tule Well (TUL; $\pi = 0.0098$) and the Pinacate lava flow (PIN and MEX; $\pi = 0.0100$ and 0.0120, respectively). The nucleotide diversity in populations from southern Arizona is two to 20 times higher than seen in any other population surveyed here. Nucleotide diversity ranged from $\pi = 0.0006$ to 0.0049 in other populations. Tajima's D was slightly negative for all populations but Carrizozo (CAR), but no values were significantly different from zero.
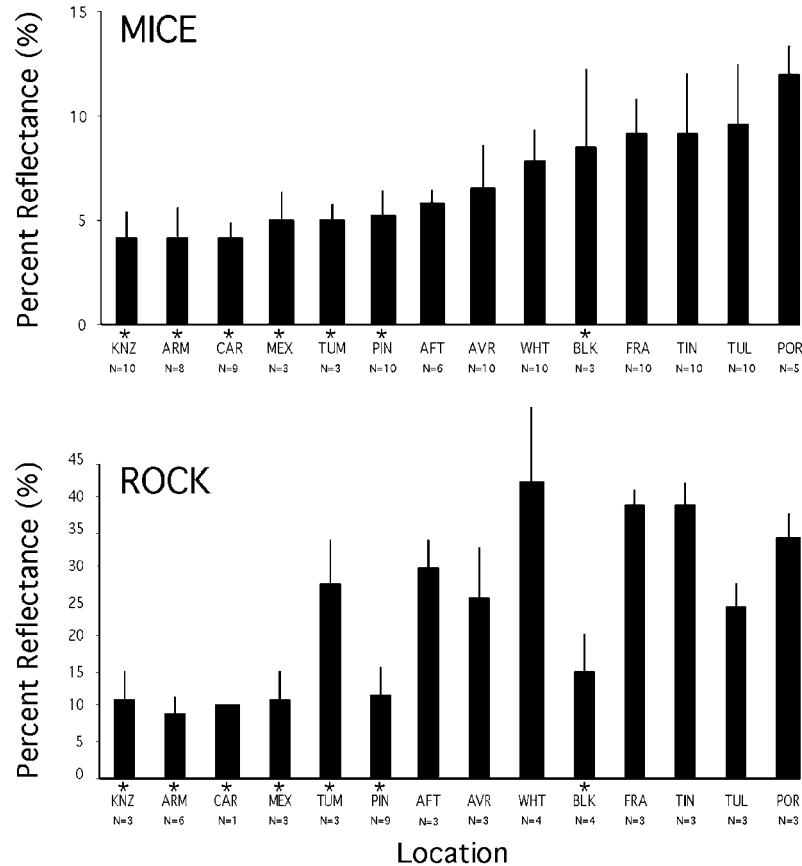
Figure 2 Spectrophotometry measurements for *C. intermedius* and their corresponding rock habitat for 14 populations. Color is measured as a percent reflectance of a white standard. Localities are aligned from darkest to lightest mice. Asterisks indicate the mice and rocks that correspond to volcanic habitats. Bars represent one standard deviation. Sample sizes are given below.

Table 2 Summary statistics of mtDNA loci by population

Population	N	bp	S[a]	Nhap[b]	π	θ	Tajima's D
All populations	107	1111	136	67	0.0143	0.0237	−1.303
Northern Arizona							
BLK*	3	1111	1	2	0.0006	0.0006	
WHT	10	1111	4	5	0.0012	0.0013	−0.340
Southern Arizona							
TIN	9	1176	23	7	0.0071	0.0072	−0.043
TUL	10	1111	38	9	0.0098	0.0121	−0.902
PIN*	10	1111	39	10	0.0100	0.0124	−0.949
MEX*	3	1111	20	3	0.0120	0.0120	
Central Arizona							
AVR	10	1176	15	5	0.0049	0.0048	−0.123
TUM*	3	1111	6	2	0.0036	0.0036	
Southern New Mexico							
POR	5	1111	6	4	0.0023	0.0026	−0.668
AFT	6	1176	3	3	0.0010	0.0011	−0.447
KNZ*	10	1176	11	6	0.0021	0.0033	−1.586
Central New Mexico							
FRA	10	1111	9	5	0.0019	0.0029	−1.443
ARM*	9	1111	12	6	0.0036	0.0040	−0.446
CAR*	9	1111	7	4	0.0028	0.0023	0.831

[a]Number of segregating nucleotide sites.
[b]Number of unique haplotypes.
Asterisk (*) indicates habitats of volcanic origin.

COIII gene

```
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000001111111111111112222222223333333333333333333344444444444444444455555555555566666666666677777
1445777899901333444556678801334467800111235566667788890111245666777788890012557778990023333567914455
5284259123842582450323436762769302362681014356735457382576712351457061450603036247694036432354509
```

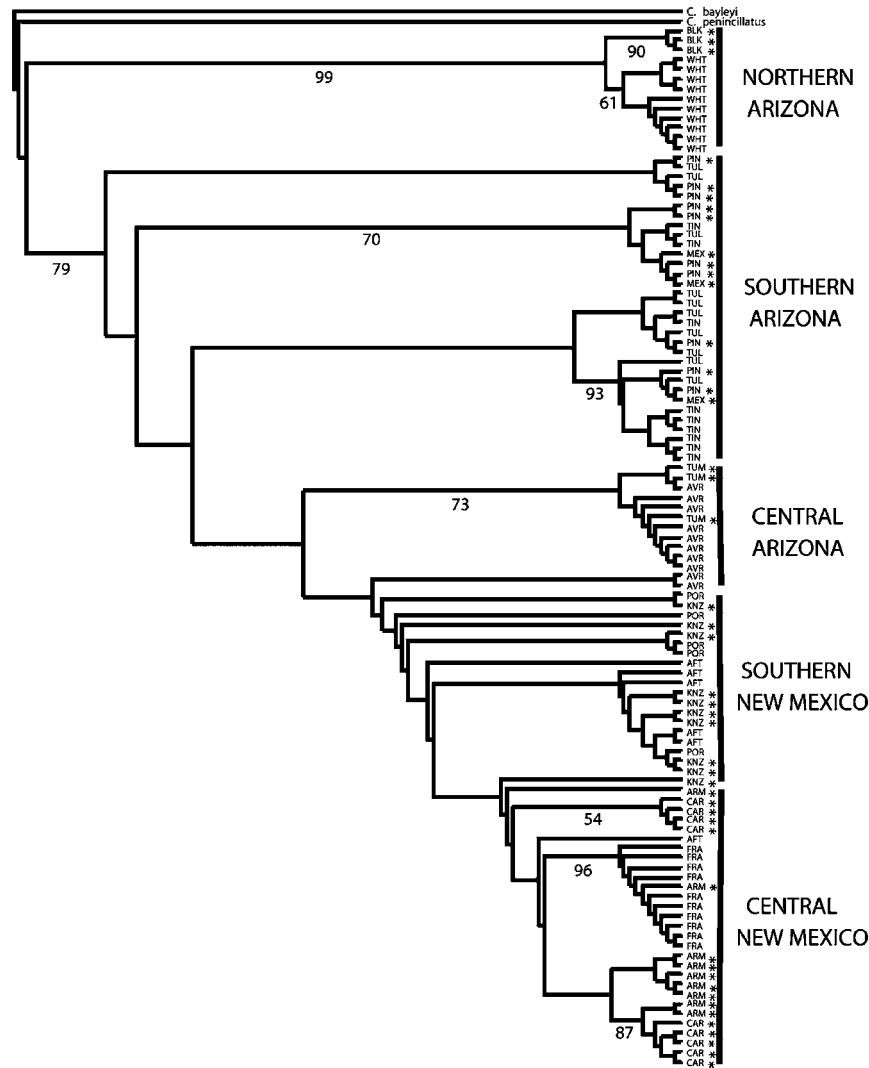| Sequence	no.	Population (no. haps)
TTAACATGTATTACCGCATTAAAGCACGAAGTGTTTCCATTTGTATCATCGTGATGTATCAGGGTAGCCCTTACATCGAATGCCTCAACAACGTTCT		
...C...A..C..T.A.................AC......C..C....C........G..G..A...TT......C.....ATT....T....C...	2	BLK(2)
...C...A..C..T.A.................AC......C..C....C........G..G..A...TT......C.....AGT....T....C...	1	
...C...A.....T.A...............G..AC......C..C....C........G..G..A...........ATT....T....C...	5	WHT(5)
...C...A.....T.A...............G..AC......C..C....C......CG..G..A...TT.....G.....AGT....T....C...	1	
...C...A.....T.A...............G..AC......C.AC....C......CG..G..A...TT..........ATT....T....C...	1	
...C...A.....T.A...............G..AC......C..C....C......CG..G..A...TT..........ATT....T....C...	1	
...C...A.....T.A...............G..AC......C..C....C......CG..G..A...TT.....G.....ATT....T....C...	1	
............T.....C..G..G.........T.G......C........C.................G................	3	TIN(7)
............T.....C.GG..G.........T.G......C..C........C.................G................	1	
............T.....C..G..C.....T.G......C..C........C.................G...........T.....	1	
............T.....C..G..G.....C....T.G......C..C......C..A.............G...........T.....	1	
............T.....G..G................C.............T.G.....TG....T.....	1	
.....G.......T.................G....CT................G.............	1	
.....G.......T.................G....CT................G...........A.....	1	
............T....................G...C.....C............C.............TG..........	2	TUL(9)
....C..................	1	
............T.....C..G.TG........T.G......C........C.............C.............	1	
...........T...C..............................C.................G...............	1	
.......A.....T.......G......C..........C.............A.......T........A.	1	
.......A.....T.......GT......C..........C.............A.......TG........A.	1	
............T................................C......T.............TG.........	1	
.....G.......T....................CT.............AG.............	1	
..G..........T.....C.....G.........CT.G......C.......C......A..................	1	
....	1	PIN(10)
....	1	
............T...C....A........C............C...............G........	1	
............T.....C..G..G.........T.G......C......C........T.................	1	
.......A.....T.......G......C..........C.............A.......TG........A.	1	
.....G.......T....................C...........................C.........	1	
.....G.......T....................C..............G......C.........	1	
.....G.....GT....C..G.........CT................	1	
.....G.....GT....C...........C.....CT................	1	
............T.....C..G..G.........T.G......C..A.....C..A...A................	1	MEX(3)
............T.....C..G..G.........T.G......C....G....C................	1	
.....G.....GT....C.............C.....CT..............	1	
.....G.......T...........A..........C.....GCT......C................	1	
..........T.........A..........T.............C...........G............C...	5	AVR(5)
..G......G...T.........A.....T......C........A.............C...	2	
......A.....T.........A..........T.............C...............C...	1	
..........T.........A..........T.............C...........C...	1	
.C..........T.........A..........T.............C........A.................	1	
.C..........T.........A..........T.............C........A.................	2	TUM(2)
..........T.........A..........T.............C...............C...	1	
..G..........T.......G.........T.............C.......A...............	2	POR(4)
..G..........T.................T.......C.A.......A..............	1	
..G..........T.................T.............C.......A...............	1	
..G..........T.................T.............C.......A.............G.....C	1	
..G..........T.................T.............C.......A.................	3	AFT(3)
..G..........T.................T.............C.......A......G............	2	
..G......G...T...........A.....T.............C.......A.............	1	KNZ(6)
..G..........T.................T.............C.......A............	4	
..G..........T.......G.........T.............C.......A..........C............	1	
..G..........T.................T.............C.......A.............	1	
..G..........T.................T.............C.......A.............	1	
..G..........T..G..............T.............C.......A.............	1	
..........T.......T.........T.............C.....A.....A..C.............G...	1	FRA(5)
..G......G...T...........A.....T.............C.......A...G.....T..............T.....	6	
..G......G...T...........A.....T.......A.C......A......T...................	1	
..G.T....G...T...........A.....T.............C.......A...G............T.....	1	
..G......G...T...........A.....T.............C.......A...G............T.....	1	
..G......G...T...........A.....T.......GC......A...G............T.....	1	
..G.....CG...T...........A.....T.............C..C......A............	3	ARM(6)
..G.....CG...T...........A.....T.............C..C......A............	2	
..G......G...T...........A.....T.............C.......A.....C...C............	1	
..G..........T...........A.....T.............C.......A.............	1	
..G......G...T...........A.....T.............C.......A......C.........	1	
..G......G...T...........A.....T.............C.......A...G......T.............T.....	1	
..G.....C.T...........A.....T.............C.......A.............	4	CAR(4)
..G......G...T...........A.....T.............C.......A.............	2	
..G......G...T...........A.....T.............C.......A.............	2	
..G......G...T.A.........A.....T.............C.......A.............	1	
```

**Figure 3** Aligned haplotypes and polymorphic sites. The positions of the sites are indicated on the top of the table. For each site, the consensus nucleotide is given; dots indicate identity to the consensus. The 71 haplotypes are arranged by frequency within each population, which is given on the right. Haplotypes are unique to collecting locales with four exceptions: shared alleles in PIN-TUL, KNZ-AFT and ARM-FRA. Site 816 is the only polymorphic site in the intergenic region and was not sequenced in all individuals.

## ND3 gene

```
                                      no. individuals
0 00000000000000000001111111111111111111
8 88888899999999999990000000000011111111
1 66789902233334589992234566779000122335
6 179479546245919828958408170680275141467    Population
                                              (no. haps)
  CAATGACATGTAACAGATCGTTTACAGTCTTTAGGACAT

C ......T.CA....G.GCT.....T.TCTC...AA....  2  BLK(2)
C ......T.CA....G.GCT.....T.TCTC...AA....  1
C ......T.CA....G.GCT.....T.TCTC...AAG...  6  WHT(5)
C ......T.CA....G.GCT.....T.TCTC...AAG...  1
C ......T.CA....G.GCT.....T.TCTC...AAG...  1
C ......T.CA....G.GCT.....T.TCTC...AAG...  1
C ......T.CA....G.GCT.....T.TCTC...AAG...  1
T .............G..C.......G.........G...  3  TIN(7)
T .............G..C.......G.........G...  1
T .............G..C.......G.........G...  1
T .............G..C...............G...  1
T .....G.......G..C...............G...  1
T .....G.......G..C....C...........G...  1
T .....G.......G..C...............G...  1
...........G..C.......G...........G...  2  TUL(9)
.................C...............G...  1
...........G..C.......G...........G...  1
T..C...........C...C...............G...  1
...........G..C...............G...  1
.G.........G..C...............G...  1
...........G..C...............G...  1
.....G.......G..C...C...........G...  1
...........G..C...............G...  1
.................................G...  1  PIN(10)
.................C...............G...  1
T..C...........C...C...........G..C  1
...........G..C.......G...........G...  1
...........G..C...............G...  1
.....G.......G..C...............G...  1
.....G.......G..C...............G...  1
.....G.......G..C...............G...  1
.....G.......G..C...............G...  1
...........G..C.......G...........G...  1
...........G..C.......G...........G...  1  MEX(3)
.....G.......TG..C...............G...  1
.....G.......GA..C...............G...  1
...C.........G.G..C...........C....G...  5  AVR(5)
.................C...............GT..  2
...C.........G..C...........C....G...  1
...C.........G..C...........C....G...  1
...........G..C...............C..G...  1
.................G..C...........G...  2  TUM(2)
T...C.......G..G..C...........C....G...  1
...........G..C...............G...  2  POR(4)
.....G.......G..C...............G.G.  1
T .............G..C...............G...  1
T .............G..C...............G...  1
T .............G..C...............G...  3  AFT(3)
T .............G..C...............G...  2
T .............G..C...............G...  1
T .............G..C...............G...  4  KNZ(6)
T .............G..C...............G...  1
T ..G..........G..C.....G...........G...  1
T ..G..........G..C...............G...  1
T .............G..C...............G...  1
T .............G..C...............G.G.  1
T ....C.T......G..C...............G...  6  FRA(5)
T .....T.......G..C...............G...  1
T ....C.T......G..C...............G...  1
T ....C.T......G..C...............T..G...  1
T ....C.T......G..C...............G...  1
T .....T...C...G..C...............G...  3  ARM(6)
T .....T...C...G..C..A.............G...  2
T .....T.......G..C...............G...  1
T .....T.......G..C...............G...  1
T .....T.......G..C...............G...  1
T ....C.T......G..C...............G...  1
......T.......G..C.A.............G...  4  CAR(4)
......T.......G..C...............AG..  2
......T.......G..C...............AAG..  1
...C..T.......G..C...............G...  1
```

Figure 3 Continued

Phylogenetic relationships: Phylogenies were constructed using the HKY85 + Γ model of nucleotide evolution and the estimates of ti/tv = 6.05 and γ = 0.21 from Modeltest. The topologies were rooted with the sister species *C. penicillatus* and *C. bayleyi*. Here, we present the NJ topology (Figure 4). Parsimony and ML trees showed similar topologies and shared the following features: the two northern Arizona populations, BLK and WHT, were reciprocally monophyletic and together formed the sister group to all other populations, suggesting that *C. intermedius* may have originated in northern Arizona. On the other hand, all mice from New Mexico group together as a recently derived clade. Within New Mexico, however, none of the individual populations formed monophyletic groups. Owing to the derived position of these populations, this pattern may be due to incomplete lineage sorting. Alternatively, continuous gene flow among New Mexico populations could also produce this pattern. Finally, individuals inhabiting lava flows were scattered across the topology.

The multidimensional scaling (MDS) plot (Figure 5), based on pairwise F_{ST} measures between populations, displayed distinct patterns of population clustering and was very similar to the clustering observed with the phylogenetic tree (Figure 4). Each of the five geographic regions was highly clustered, and there was a clear demarcation between Sonoran and Chihuahuan desert populations. Both the phylogenetic tree and the MDS plot separated the northern Arizona populations (BLK + WHT) from all other populations. In addition, populations from all five geographic regions clustered independently. Particularly strong clustering was observed among the three populations within southern New Mexico and among the four populations within southern Arizona. The correlation between the genetic distance (as measured by F_{ST}) and the two-dimensional plot was high ($r = 0.94$).

Population structure: Haplotype diversity was significantly partitioned among populations and geographic regions in *C. intermedius*. In the one-factor AMOVA of all populations, most of the variation was explained by differences among populations ($\Phi_{ST} = 0.69$). Additional AMOVA analyses revealed that partitioned variation between the two biogeographic regions, the Sonoran and Chihuahuan deserts, was significant ($\Phi_{CT} = 0.21$, $P < 0.004$), but that most of the variation was partitioned among populations ($\Phi_{ST} = 0.65$, $P < 10^{-5}$). Variation partitioned among the five geographic regions was high ($\Phi_{CT} = 0.64$, $P < 10^{-5}$), as was the variation partitioned between the two northern Arizona populations (WHT + BLK) compared to all other populations ($\Phi_{CT} = 0.34$, $P < 10^{-5}$).

Correlation between phenotype, genotype and geography

Using Mantel tests, we partitioned the data among genetic, geographic and phenotypic components (Table 3). There is a significant correlation between genetic (mtDNA) variation and geographic distance, where geographic distance explains 40% of genetic variation ($P = 0.002$). In addition, geographic distance was correlated with phenotype ($r = 0.24$), although the correlation between geography and genetic variation was stronger. However, there is no correlation between genetic variation and phenotype ($r = 0.08$, $P = 0.29$), and when we control for the effects of geography, the correlation is further weakened ($r = 0.00$, $P = 0.48$).

Figure 4 NJ phylogeny of *C. intermedius* populations. Topology is rooted with *C. penicillatus* and *C. baileyi*. Asterisks indicate individuals inhabiting lava. Geographic regions are indicated on the right. Bootstrap support is indicated under the internal branches.

Figure 5 MDS plot of the 14 populations based on a matrix of pairwise F_{ST} measures. Abbreviations for populations follow those given in Table 1. Diagonal line separates Arizona Sonoran desert populations and New Mexico Chihuahuan desert populations.

Population divergence and local adaptation

To test the role of local adaptation in phenotypic evolution, we examined the extent of gene flow occurring between lava populations and the nearest population inhabiting light-colored rocks. In this study, there were four pairs of populations for which light and melanic mice were found in close geographic proximity (Table 4). We found substantial gene flow in several of these comparisons. The population migration rate was highest between PIN and TUL ($Nm > 60$) and also high in KNZ and AFT ($Nm = 8$). Both of these estimates of gene flow are higher than those that typically lead to population differentiation. This pattern can also be observed in the phylogenetic tree (Figure 4), where individuals from these populations are intermingled along the tips. However, surprisingly there was some population structure between ARM and FRA as well as TUM and AVR ($Nm = 5$ and 2, respectively), suggesting that these lava populations are at least partially isolated from nearby neighboring light-colored populations. In addition, we used a MCMC method to jointly estimate

gene flow and time of divergence in order to determine if shared polymorphisms are due to recurrent gene flow or simply recent ancestry of populations (Table 5). In all four pairwise comparisons, we were able to reject the hypothesis of no recurrent gene flow ($M = 0$). In all cases, estimates of gene flow were lower when recent ancestry was considered (two- to 10-fold), but the rank order based on Nm remained the same.

In addition, we estimated levels of gene flow connecting lava-dwelling populations to examine the hypothesis that melanic mice evolved once in New Mexico and migrated from one lava flow to a neighboring (perhaps younger) lava flow (Table 4). We found the level of gene flow was similar between CAR and KNZ compared to ARM and KNZ ($Nm \sim 0.5$). However, migration rates were higher between the geographically closer populations of CAR and ARM ($Nm = 1.2$). Using the MCMC method, we found similar patterns to those revealed by F_{ST}-based statistics (Table 5). Migration rates were lower between KNZ and both ARM ($Nm = 0.14$) and CAR ($Nm = 0.02$). Migration rates were higher between CAR and ARM ($Nm = 0.40$), and we were able to reject a model of recent divergence and no gene flow ($M = 0$) between these two populations. All MCMC estimates of gene flow were lower than those based on F_{ST}. However, we noted that 95% credibility estimates are very large and encompass F_{ST} estimates of migration rate.

Discussion

We documented substantial phenotypic variation across the range of *C. intermedius*. This color variation was not, however, correlated with phylogeny, suggesting that history is not responsible for the present distribution of phenotypic variation. Although there is substantial genetic structure between geographic regions, high levels of gene flow (or recent ancestry) connect populations within a region. Despite these high levels of local gene flow, color variation is strongly correlated with habitat color in most populations, suggesting that natural selection for substrate matching is strong in this species.

Table 3 Mantel and partial Mantel test for significant correlation between genetic, geographic and phenotypic variation in 14 populations

Matrix comparison	r	P-value
MtDNA–Geography	0.40	0.002
MtDNA–Phenotype	0.08	0.285
Geography–Phenotype	0.24	0.030
MtDNA–Geography_(Phenotype)	0.40	0.001
MtDNA–Phenotype_(Geography)	0.00	0.476

Genetic distances between populations were measured using pairwise F_{ST}. Geographic distances were based on the central collecting site for each population. Phenotypic distance (Manhattan distance) was measured using the program NSYS. Significance values were determined by comparing the observed z-statistic and the expected z-statistic, generated by a randomized distribution from 1000 permutations. Parentheses indicate which factor is removed from the correlation.

Table 4 Pairwise estimates of gene flow between the four paired dark and light rock populations and separately the three New Mexico lava populations

Population comparison	F_{ST}	Nm	Distance[a] (km)
Lava population and neighboring light rock population			
PIN–TUL	0.008	60.43	16
TUM–AVR	0.172	2.41	28
KNZ–AFT	0.057	8.35	7
ARM–FRA	0.098	4.61	7
New Mexico lava populations harboring melanic mice			
CAR–ARM	0.296	1.19	118
KNZ–CAR	0.510	0.49	222
KNZ–ARM	0.517	0.47	158

[a]Approximate geographic distances, based on the center of each collecting area, are given for each pairwise comparison.

Table 5 Pairwise estimates of female effective population size ($\theta = 2N_f u$), female migration rate ($M = N_f m$) and divergence time ($T = t/N_f$), based on analysis of two mtDNA genes (COIII and ND3) using the program MDIV (Nielsen and Wakeley, 2001)

Population comparison	θ	M	T
Lava population and neighboring light rock population			
PIN–TUL	13.49 (9.38–23.09)	8.12 (3.24–19.08)	0.04
TUM–AVR	2.36 (1.53–6.25)	1.38 (0.58–9.14)	0.32
KNZ–AFT	2.65 (0.96–5.78)	1.50 (0.68–9.42)	0.26
ARM–FRA	3.04 (1.85–6.28)	1.46 (0.08–6.58)	0.98
New Mexico lava populations harboring melanic mice			
CAR–ARM	3.65 (2.12–7.26)	0.40 (0.04–6.06)	0.60
KNZ–CAR	3.16 (1.99–6.85)	0.02 (0.00–1.16)	1.30
KNZ–ARM	4.61 (2.09–9.48)	0.14 (0.04–1.84)	1.10

A uniform prior (0,10) was assumed for θ, M and T. Markov chain was simulated for 5×10^6 generations, where 5×10^5 generations were used for a burn-in time. For θ and M, values with the highest likelihood scores are given, followed by 95% credibility intervals within parentheses.

Adaptive phenotypic variation

Spectrophotometric measurements reveal a strong correlation between substrate color and the dorsal pelage of rock pocket mice ($R^2 = 0.43$). Rock substrate from lava flows show significantly lower reflectance than rocks from other areas. The dorsal pelage of mice inhabiting these lava flows has correspondingly lower reflectance. One distinct exception is the population of Black Tank (BLK) in northern Arizona in which the rock has similar reflectance to other lava flows but the mice are similar in reflectance to other nonlava mice. This anomaly likely reflects the fact that the BLK lava flow is an extension of the relatively young Sunset Crater (<800 years old); consequently, there may have been insufficient time for melanic mice to evolve in this population. In addition, there are no populations of melanic mice nearby from which melanic migrants may invade.

In other cases, migration from neighboring populations may be responsible for the presence of melanic (or dark) phenotypes in some areas. For example, the Tumamoc Hill population, which occurs on olivine basalt, is separated by just a few miles from a melanic population of mice (Dice and Blossom, 1937), which occurs on the extremely dark basalt of Black Mountain, which we were unable to sample. Owing to their close geographic proximity, it is likely that gene flow occurs between these two populations. In fact, Blossom (1931) described a dark-colored race 'nigrimontis' from Black Mountain and Tumamoc Hill. Therefore, Black Mountain may be the source of the Tumamoc Hill mice and may explain why the Tumamoc mice are significantly darker than their habitat. Using genetic data, we also explored migration of melanic mice among the lava-dwelling New Mexico populations (see below).

Genetic variation

In the northern range of *C. intermedius*, we found a high level of genetic partitioning between the five geographic regions ($\Phi_{CT} = 0.64$). This pattern is also reflected in the phylogenetic tree, which shows a general grouping of individuals by geographical region, although some of these groups are not monophyletic (Figure 4). The MDS plot, based on pairwise F_{ST} estimates, also reveals strong clustering by geographic region (Figure 5). As rock pocket mice exclusively inhabit rocky areas and are replaced by sand-dwelling species (eg *C. penicillatus*) in nonrocky areas, their habitat is largely discontinuous throughout its range, which may underlie the population structure observed at this geographic scale. At a smaller scale, however, substantial gene flow sometimes occurs between neighboring populations (Table 4).

It is important to note that the genetic analyses we performed are based on a single molecular marker, mtDNA. In addition, mtDNA tracks only female migration. However, because there is no evidence for sex-biased dispersal in heteromyid rodents, mtDNA may be an accurate predictor of the average gene flow for both sexes (Jones, 1993).

Biogeographical history of *C. intermedius*:
Phylogenetic analysis reveals a strong split between the two northern populations (BLK and WHT) and the rest of the populations. This northern clade is basal in the phylogeny. AMOVA analysis suggests that 34% of genetic variation was explained by differences between the northern populations and the remaining populations. These results suggest that *C. intermedius* may have originated in northern Arizona and expanded to southern Arizona and eventually east into southern and central New Mexico. The average pairwise divergence between these two northern Arizona populations and the remaining populations was 0.035. Assuming a 2% per million year clock, *C. intermedius* may have expanded from northern Arizona southward about 1.5 million years ago. Interestingly, this corresponds to the age of the Pinacate lava flow, suggesting that pocket mice were likely present in the southern Arizona region during the formation of the Pinacate lava flow.

Correlation between phenotype, genotype and geography

Are morphological patterns reflected in the genetic structure of *C. intermedius*? In this case, removing the geographic component of the phenotypic variation among populations resulted in no correlation between phylogeny and color variation ($r = 0.00$). The phenotypic patterns in *C. intermedius* may reflect processes operating on a spatial or temporal scale much smaller than that reflected in broad-scale geographic patterns, resulting in a poor correlation between color patterns and geography and phylogeny. This result suggests that adaptation to local environments (ie natural selection) is a key force driving morphological diversity in this system and that historical contingency plays a relatively small role.

Local adaptation

To determine the amount of gene flow occurring between populations that differed in their substrate color (and phenotype), we calculated gene flow in four comparisons between geographically proximate populations inhabiting lava and light rock. We found that in each case substantial gene flow occurs over short distances (Nm ranged from 60.5 to 2.5 using F_{ST}-based statistics and 8.12–1.46 using MCMC estimates), suggesting that natural selection must be strong in order to maintain habitat-specific color patterns. In the Pinacate region, Hoekstra *et al* (2004) estimated selection coefficients against light mice inhabiting lava as high as 0.39. It is important to recognize, however, that our estimates of gene flow from F_{ST} are based on a model of migration–drift balance at equilibrium. Inherent in these calculations are a number of assumptions which may not be biologically realistic (Whitlock and McCauley, 1999).

Evolution of melanism in New Mexico populations:
Molecular analysis of pigmentation genes shows that melanic pocket mouse populations have evolved independently in the Pinacate population of Arizona and the lava populations of New Mexico (Hoekstra and Nachman, 2003). However, it remains unclear if melanic populations have evolved repeatedly within New Mexico, that is, independently in Armendaris (ARM), Carrizozo (CAR) and Kenzin (KNZ) populations. Our analysis suggests that substantial gene flow may occur among these lava-dwelling pocket mouse populations in New Mexico, raising the possibility that melanic mice evolved once in New Mexico and migrated to other lava populations. Again, however, we caution that the assumption of migration–drift equilibrium is

unlikely to be met, especially among the New Mexico populations. These populations are relatively young (see below) and at least some of the shared variation among populations may reflect ancestral polymorphism for which lineage sorting is incomplete, thus leading to an overestimate of the true level of gene flow. The MCMC method of Nielsen and Wakely (2001), however, assesses the relative roles of migration and isolation as causes of the observed differentiation between populations.

We estimated the approximate time when pocket mice invaded New Mexico from southern Arizona in order to identify the maximum time for adaptation to local conditions to occur in New Mexico. The average pairwise difference between the New Mexico populations and the central Arizona populations was 0.0095. Assuming a molecular clock of 2% mtDNA divergence per million years, these mice first appeared in New Mexico roughly 500 000 years ago. The ages of both the Kenzin ($\sim$500 000 years old) and Armendaris ($\sim$750 000 years old) lava flows are similar to or older than this estimate. Thus, the maximum time for local adaptation on these lava flows may be limited by the immigration of C. intermedius to New Mexico approximately 500 000 years ago.

One lava flow in New Mexico, Carrizozo, is less than 1000 years old, yet harbors a population of mice which have uniformly melanic dorsal pelage (Dice and Blossom, 1937). Given the young age of the lava flow, this is somewhat surprising and leads us to consider the relative probabilities that these melanic mice arose from a new mutation or from migration of a new allele. We can get a rough estimate of these probabilities as follows. Mutation rates are likely to be approximately 10^{-5} per locus per generation (eg Schlager and Dickie, 1971), corresponding to about 10^{-8} per nucleotide site per generation. If there are 100 sites in the genome at which mutations can produce dark color, the overall mutation rate to dark color is about 10^{-6} per genome per generation. If we assume that the population size is approximately 10^4 and that 10^3 generations have elapsed since the appearance of the lava flow (one generation per year), then there has been sufficient time for approximately 10 darkening mutations to appear. In contrast, if the number of migrants (Nm) from the nearest dark population (Armendaris) is approximately one per generation (F_{ST}-based estimate; Table 4) or one individual every two generations (MCMC-based estimate; Table 5), then there has been sufficient time for approximately 500–1000 dark alleles to be introduced by migration. These rough calculations suggest that migration is a more likely source of dark alleles in the Carrizozo population than is mutation.

Migration rates are likely to be high between Armendaris and Carrizozo because there is suitable rocky habitat (the San Andres Mts), which spans much of the area between these lava flows. Additionally, melanic mice may also have migrated from the Kenzin lava (although at a lower rate), increasing the likelihood of introducing melanic alleles via migration. While it is more likely that melanic migrants are responsible for the melanic population at Carrizozo versus new mutations in the Carrizozo population, identifying the genetic basis of melanism in these populations will directly address this question.

Together, our results suggest that strong selection is likely to be the driving force in promoting morphological diversity in this system. We observe high levels of gene flow in several neighboring populations that differ in substrate color, and the observation that mice closely match their substrate despite this gene flow indicates that strong selection is maintaining habitat-specific phenotypes. Migration is commonly thought to be a homogenizing force impeding local adaptation; however, here we raise the possibility that migration between lava flows may also promote local adaptation by introducing beneficial alleles into neighboring populations experiencing similar selective regimes.

Acknowledgements

References

Benson SB (1933). Concealing coloration among some desert rodents of the southwestern United States. Univ Calif Publ Zool 40: 1–69.

Blossom PM (1931). Relation between color of desert rodents and of the soil. Carnegie Inst Washington Yearb 30: 266.

Castellano S, Balletto E (2002). Is the partial Mantel test inadequate? Evolution 56: 1871–1873.

Dice L, Blossom PM (1937). Studies of mammalian ecology in southwestern North America, with special attention to the colors of desert mammals. Publ Carnegie Inst Washington 485: 1–25.

Excoffier L, Smouse PE, Quattro JM (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131: 479–491.

Haldane JBS (1948). The theory of a cline. J Genet 48: 277–284.

Hoekstra HE, Drumm KE, Nachman MW (2004). Ecological genetics of coat color variation in pocket mice: geographic variation in selected and neutral genes. Evolution 58: 1329–1341.

Hoekstra HE, Nachman MW (2003). Different genes underlie adaptive melanism in different populations of rock pocket mice. Mol Ecol 12: 1185–1194.

Jones T (1993). The social systems of heteromyid rodents. In: Genoways HH, Brown JH (eds) Biology of the Heteromyidae. American Society of Mammalogists, pp 575–595.

Lenormand T (2002). Gene flow and the limits to natural selection. Trends Ecol Evol 17: 183–189.

Manly BFJ (1986). Randomization and regression methods for testing association with geographical, environmental and biological distances between populations. Res Popul Ecol 28: 201–218.

Nachman MW, Boyer SN, Searle JB, Aquadro CF (1994). Mitochondrial DNA variation and the evolution of Robertsonian chromosomal races of house mice, Mus domesticus. Genetics 136: 1105–1120.

Nachman MW, Hoekstra HE, D'Agostino S (2003). The genetic basis of adaptive melanism in pocket mice. *Proc Natl Acad Sci USA* **100**: 5268–5273.

Nei M, Li WH (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc Natl Acad Sci USA* **76**: 5269–5273.

Nielsen R, Wakeley J (2001). Distinguishing migration from isolation: a Markov chain Monte Carlo approach. *Genetics* **158**: 885–896.

Palsboll PJ, Berube M, Aguilar A, Notarbartolo-Di-Sciara G, Nielsen R (2004). Discerning between recurrent gene flow and recent divergence under a finite-site mutation model applied to North Atlantic and Mediterranean sea fin whale (*Balaenoptera physalus*) populations. *Evolution* **58**: 670–675.

Posada D, Crandall KA (1998). MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**: 817–818.

Raufaste N, Rousset F (2001). Are partial Mantel tests adequate? *Evolution* **55**: 1703–1705.

Rousset F (1997). Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics* **145**: 1219–1228.

Rousset F (2002). Partial Mantel tests: reply to Castellano and Balletto. *Evolution* **56**: 1874–1875.

Rozas J, Rozas R (1999). DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics* **15**: 174–175.

Schlager G, Dickie M (1971). Natural mutation rates in the house mouse. Estimates for five specific loci and dominant mutations. *Mutat Res* **11**: 89–96.

Schneider S, Roessli D, Excoffier L (2000). *Arlequin: A Software Program for Population Genetics Data Analysis*. Genetics and Biometry Lab, Department of Anthropology, University of Geneva.

Slatkin M (1985). Gene flow in natural populations. *Annu Rev Ecol Syst* **16**: 393–430.

Slatkin M (1993). Isolation by distance in equilibrium and non-equilibrium populations. *Evolution* **43**: 264–279.

Swofford DL (1999). *Phylogenetic Analysis using Parsimony (PAUP)*. Sinauer: Sunderland, MA.

Tajima F (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**: 585–595.

Travisano M, Mongold J, Bennett A, Lenski R (1995). Experimental tests of the roles of adaptation, chance, and history in evolution. *Science* **267**: 87–90.

Waterson GA (1975). On the number of segregating sites in genetical models without recombination. *Theoret Popul Biol* **7**: 256–276.

Weckerly WF (1983). *Geographic Variation of the Rock Pocket Mouse, Perognathus intermedius Merriam, on the Perdro Armendariz Lava Flow of South-Central New Mexico*. Master's Thesis. Eastern New Mexico University, Portales, NM.

Whitlock M, McCauley D (1999). Indirect measures of gene flow and migration: *F*-ST not equal 1/(4*Nm*+1). *Heredity* **82**: 117–125.

Wilson AC, Cann RL, Carr SM, George M, Gyllensten UB, Helmbychowski KM *et al* (1985). Mitochondrial DNA and two perspectives on evolutionary genetics. *Biol J Linn Soc* **26**: 375–400.

Appendix A1

Previously published mtDNA sequences from three sites (Hoekstra *et al*, 2004) are as follows: PIN – AY648411, AY648511, AY648514-9, AY648525-6; TUL – AY648417-8, AY648420-1, AY648527-8, AY648533, AY648537, AY648544, AY648552; MEX – AY648479-80, AY648503.